



OPEN

Angry facial expressions elicit a late attentional withdrawal

Daniela Ballotta^{1,5}✉, Riccardo Maramotti^{1,2,3,5}, Eleonora Borelli⁴, Fausta Lui¹ & Giuseppe Pagnoni¹

Facial expressions of emotion are salient social cues that can trigger adaptive behavioral responses—such as facilitated approach or avoidance movements in response to positive or negative stimuli, respectively. We conducted a behavioral experiment testing the hypothesis that while both happy and angry faces initially capture attention, angry expressions subsequently prompt a withdrawal of attention—a pattern consistent with aversive behavioral tendencies. Using a target-detection paradigm, we varied the spatial location and timing (50 ms vs. 350 ms) of targets relative to face stimuli. Results supported the hypothesis: angry faces led to faster responses to targets appearing nearer to the observer (i.e., farther from the face) at 350 ms, suggesting attentional retraction. A follow-up fMRI experiment replicated this design with only the 350 ms delay. Behaviorally, responses were faster to targets near happy faces and away from angry ones. These effects correlated with subjective reports of being attracted or repelled by the expressions, as well as with individual differences in empathy and trait avoidance. Neuroimaging data revealed an interaction between emotion and target location in early visual cortices, modulated by subjective attraction-retraction ratings. These findings support the idea of attention disengaging from angry faces, with both behavioral and neural correlates linked to individual affective dispositions.

The relationship between emotion and attention has long been debated in the psychological and neuroscientific literature (for reviews, see^{1–7}). Since our cognitive resources are limited, different pieces of visual information compete for processing, and a prioritization of salient stimuli is necessary. Selective attention allows us to focus on relevant stimuli while ignoring irrelevant information^{8–11}, and emotional significance is one of those aspects that flag a stimulus or an event for priority processing¹.

Converging evidence strongly suggests that biologically and socially relevant stimuli, such as facial expressions of emotion, are perceptually salient in the human environment, and they seem to be processed in a rapid and automated way¹². The high temporal resolution of event-related potentials (ERP) has been used to investigate the temporal dynamics of emotional face processing and of the interplay between emotion and attention⁶. According to the three-stage model of facial expression processing¹³, detection of negatively-valenced stimuli occurs during the first stage, other emotional expressions are distinguished from neutral ones during the second stage, and a finer differentiation of various emotional facial expressions takes place only in the third and final stage. In EEG recordings, an early visual evoked potential component, occurring 80–90 ms after stimulus onset (P100), seems to be modulated by global emotional content even when emotional information is task-irrelevant^{14,15}, whereas the emotional modulation of later components, such as the N170^{14,16,17} and the P300, depends on the availability of attentional resources^{13,18}. Thus, it has been argued that beyond the initial automatic detection of emotion, spatial attention can modulate later stages of facial expression processing¹⁹. Conversely, emotional stimuli can influence the allocation of attention. For example, a specific influence of fearful facial expression on spatial selective attention was demonstrated by Ellena et al.²⁰, although, more generally, the affect-specific spatial pattern and dynamics of selective attention shifts remains yet to be fully clarified.

A recent fMRI study by our group found a significantly different response of the early visual areas to negative and positive valence of facial expressions, during an implicit task²¹, possibly through a facilitation of the activity of the amygdala by positive faces, and of the fusiform face area by negative faces. Such findings support the hypothesis that the neural correlates of emotional expression detection differ at early ‘sensory-specific’ processing phases^{14,15}. Furthermore, the activity of sensory cortices is known to be modulated by both emotional and

¹Department of Biomedical, Metabolic and Neural Sciences, University of Modena and Reggio Emilia, 41125 Modena, Italy. ²Department of Physics, Informatics and Mathematics, University of Modena and Reggio Emilia, 41125 Modena, Italy. ³Department of Mathematics and Computer Sciences, University of Ferrara, 44121 Ferrara, Italy. ⁴Department of Medical and Surgical, Maternal-Infantile and Adult Sciences, University of Modena and Reggio Emilia, 41125 Modena, Italy. ⁵Daniela Ballotta and Riccardo Maramotti contributed equally to this work. ✉email: daniela.ballotta@unimore.it

spatial attention^{22–24}, where the influence of emotional content is mediated by automatic bottom-up processes, while fronto-parietal attentional networks²⁵ exert a top-down influence^{23,26} in allocating spatial attention to emotional, especially threatening, facial expressions³ and in selecting visual locations^{27–29}.

The attentional bias towards emotion plays a crucial role in regulating the interaction of individuals with their conspecifics, and more generally with their environment. Facial expressions are extremely powerful in communicating emotional states and behavioral intentions of an agent^{30–32}, and can trigger different and adaptive behavioral reactions in the onlooker. In fact, a fundamental link between emotions and action tendencies such as approach and avoidance has been postulated³³, and the organization of emotions in two different motivational systems has been proposed by Lang et al.³⁴. These motivational systems may guide the individual in preparing adequate and adaptive responses when emotionally relevant stimuli are present in the environment: approach of positive-valenced stimuli would be mediated by an appetitive system, whereas the avoidance of negative stimuli would depend on a defense-oriented system. Since the seminal study of Solarz³⁵, an effect of compatibility between affect and action tendencies (‘affective compatibility effect’) has been largely replicated using motor tasks and emotional words^{36–38}, pictures^{39,40}, and facial expressions^{41,42} as stimuli, showing a facilitation of approaching movements with positive items and of avoiding movements with negative items. A direct mapping of muscle movements onto stimulus valence was suggested by Chen et al.³⁶ to explain this effect. The valenced stimuli would automatically activate components of the behavioral repertoires of corresponding motivational systems (positive stimuli → flexion actions, ‘bring the stimulus closer’ approach system); negative stimuli → extension actions, ‘push the stimulus away’ avoidance system). However, contrasting findings exist in the literature, with reports showing that both positive and negative stimuli can elicit appetitive and/or aversive behaviors^{43–46} (see also meta-analyses by^{47,48}).

The aim of the present study was to investigate the impact of emotional stimuli on attentional allocation, under the hypothesis of a compatibility effect between the direction of the attentional shift induced by emotional stimuli and the valence of the latter. More specifically, on the behavioral level, we tested the hypothesis that, even though both positive and negative faces are quick to command attention (as early as 100 ms after stimulus onset³, negative stimuli may affect attentional allocation in a paradoxical way: the initial transient phase of attentional capture by the negative stimulus may be swiftly followed by a withdrawal of the attentional spotlight away from it and towards the observer, a process directionally congruent with the behavioral pattern of an aversive reaction. This corresponds to the so-called ‘vigilance-avoidance’ hypothesis, which has also been proposed to play a role in maintaining maladaptive response patterns in anxious individuals^{49,50}.

Notably, such attentional retraction would not occur for positive stimuli, which could be associated with an opposite tendency of attentional approach—although not necessarily of the same magnitude, given the greater behavioral relevance of threatening stimuli. As a consequence, during this later phase of attentional dynamics, faster response times to target stimuli would be expected when such stimuli are presented either perceptually close to positive faces, or far from negative faces (i.e., closer to the observer). For what concerns the neural correlates of this effect, we expected an involvement of visual areas, since both the emotional valence of the stimulus and spatial attention are known to modulate the activity of visual cortices^{21,23,24}. More specifically, we hypothesized that the target-related activation in early visual cortices would be greater when the target appeared in a location close to the current locus of attention—which in turn depended by the valence of the face stimulus—compared to conditions where target location and attentional locus differed (see Fig. 1).

Experiment 1: behavioral study

To verify the effectiveness of a working choice for the delays between the face and the response target presentations and to demonstrate the bi-phasic attentional shifts during threatening facial expression processing, we conducted first a behavioral pilot study. Based on the extant literature^{3,6}, we selected two candidate delay values of 50 ms and 350 ms and tested whether they could effectively separate—based on response times—an early attentional phase, during which attention would be captured by both positive and negative face stimuli, from a later one, where attention would have already retracted from negative faces (but not from positive ones).

Methods

Participants

Forty-seven right-handed young subjects (35 females, 12 males, mean age: 23.7 ± 5.4 years; range 18–44 years) took part in the pilot study. A history of psychiatric or neurological disease and current use of any psychoactive medications were considered as exclusion criteria. The experimental protocol was approved by the local Ethics Committee (protocol number: CE 134/2014/SPER/AOUMO) and all subjects gave their written informed consent to take part in the study.

Face stimuli

Black and white photographs of emotional facial expressions of happiness ($n = 36$, positive stimuli) and anger ($n = 36$, negative stimuli) were used, with 18 expressions per condition performed by female models, and 18 by male models⁵¹. Emotional faces (as opposed to words or other types of stimuli), and specifically angry expressions, were chosen because a significantly larger affective compatibility effect has been described using this category of stimuli⁴⁸. Hair and non-facial contours were masked out and pictures were balanced for brightness, contrast values and arousal across categories.

Procedure

An event-related paradigm was employed. Each volunteer performed 4 sessions of 38 trials each, for a total of 152 trials. Each trial lasted 1500 ms, with a fixed interstimulus interval (ISI) of 1500 ms. Emotional faces were visually presented on a computer screen as appearing on the back wall of a corridor rendered in a 3D

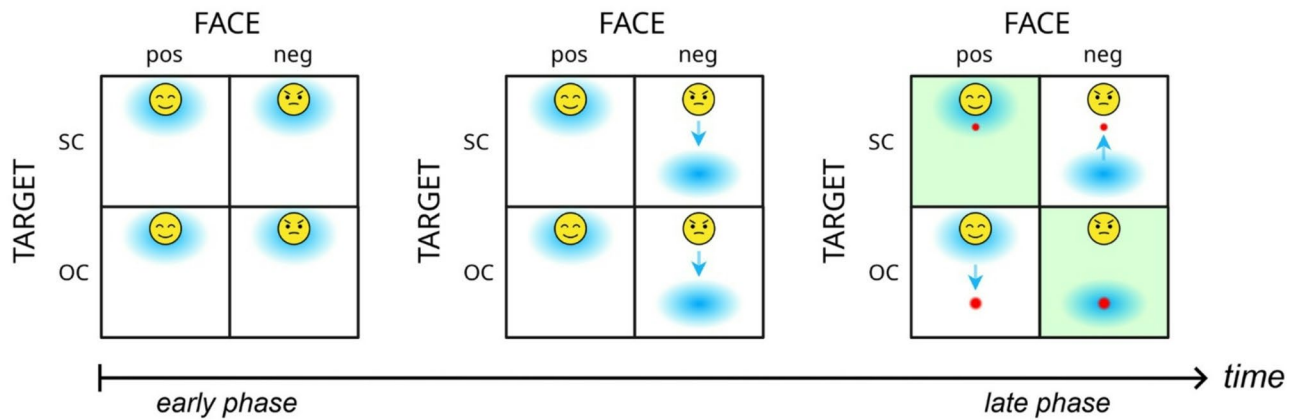


Fig. 1. Schema of the hypothesized effect of our experimental manipulation on attentional allocation, depending on the position of target appearance and the valence of the face stimulus. We assumed that, after a brief initial transient (*early phase*), attention would retract from the face stimulus for angry faces, but remain in its vicinity for happy faces (*late phase*). In the picture, the leftmost diagram represents the time of appearance of the face stimulus ($t=0$ ms), the rightmost diagram the time of target appearance, and the middle one an intermediate time when attention would shift from the vicinity of negative stimuli towards the observer—i.e., from the early to the late phase. Red dots represent target locations, blue halos indicate assumed loci of attention at each phase of the trial, and green cells in the rightmost diagram highlight conditions where target location and locus of attention coincide. The arrows indicate the direction of an attentional shift that has just occurred (middle diagram) or has to occur to respond correctly to the target (rightmost diagram). SC = stimulus-close (near the face stimulus and far from the observer), OC = observer-close (far from the face stimulus and near the observer), *pos* = positive (happy) face, *neg* = negative (angry) face.

perspective for 1500 ms. Participants were asked to respond as fast as possible to a target stimulus (a small red sphere, lasting 100 ms). A $2 \times 2 \times 2$ factorial paradigm was employed. Half of the times the red sphere appeared 50 ms after the face stimuli, and half of the times 350 ms after the face stimuli. The target was presented half of the times perceptually closer to the face stimulus (stimulus-close, SC), and half of the times perceptually closer to the observer (observer-close, OC). Half of the stimuli were facial expressions of happiness (positive faces), while the other half were angry expressions (negative faces). The volunteers were instructed to maintain the gaze on the face and respond by pressing a button (the key ‘1’ on the computer numerical keyboard) with the index finger of their right hand. To minimize the occurrence of participants responding to the initial presentation of the face stimulus instead of waiting for the target, we introduced 8 foil trials, where stimuli were presented with no corresponding target (and button press had thus to be withheld). The pilot study was conducted during the COVID pandemic, thus it was executed in a remote modality. Participants performed the task at home on their PCs, following detailed instructions sent them by email. Volunteers were asked to position themselves with the back straight and the face approximately 60 cm from the monitor. The room had to be silent, without other people or distractions of any kind. Participants were also asked to place their mobile phone away after putting it in silent mode. Stimuli were presented via E-Prime Go (<https://www.eprimego.com>), a tool offered by E-Prime 3.0 software.

Behavioral data analysis

Response times (RTs) analysis was performed using two separate Bayesian models, one for the condition with the target appearing 50 ms after the face stimulus, and the other for the condition with the target appearing 350 ms after the stimulus. Error and foil trials were excluded from the analysis. Since RT data are known to follow a non-normal, positively skewed distribution, commonly used parametric statistical tests based on the assumption of normality are generally not appropriate⁵². Although a log-normal distribution model is often able to represent RT data well⁵³, its adequacy has been questioned⁵⁴ because of the unrealistic lack of a lower boundary for RTs in the log-normal distribution. To overcome this limitation, we used a shifted log-normal distribution, which increases the applicability of the log-normal distribution by including an additional parameter (c) that shifts the distribution away from zero, setting a lower boundary for the RTs^{55,56}. We modeled the effects of valence and target location on response times using the distribution $L \sim \text{LogNormal}(\mu, \sigma^2) + c$. The scale (σ) and shift (c , the estimated shortest possible RT) parameters were considered fixed across all experimental conditions, while the median RT (μ) was estimated separately for each of the four conditions of interest. More precisely, introducing the discrete variables v and t :

$$v = \begin{cases} 0 & \text{if valence is negative} \\ 1 & \text{if valence is positive} \end{cases} \quad t = \begin{cases} 0 & \text{if target location is SC} \\ 1 & \text{if target location is OC} \end{cases}$$

the median RT in milliseconds can be expressed as $\mu = c + \exp(b_0 + vb_1 + tb_2 + tvb_3)$ (see Supplementary Materials, for an explicit derivation of the estimated median RTs for each of the four experimental conditions

from the model parameters). The parameters b_0 , b_1 , b_2 , b_3 , σ , and c were estimated via a Bayesian non-linear model, implemented in R Statistical Software (v4.3.1; R Core Team 2021) using the `brm()` function of the package `brms`⁵⁷. Results were generated by 5 Monte Carlo Markov chains (MCMC), each of them consisting of 2000 burn-in iterations and 5000 sampling iterations.

Posterior distributions of the effects of interest were evaluated using the `hypothesis()` function in `brms`, which computes effect sizes (e.s.) and 95% credible intervals (CI) for linear combinations of the model parameters. In particular, we tested the interaction between valence and target location, defined as $(OC - SC)|_{\text{pos}} - (OC - SC)|_{\text{neg}}$, as well as the simple effects of target location within positive and negative facial stimuli. Importantly, we remind the reader that whenever we report effect sizes and relative credible intervals for the Bayesian analyses throughout the paper, these refer to the corresponding contrasts of parameter estimates (and thus are adimensional quantities, not, e.g., msec for response times).

Results

On average, participants reached an accuracy of 95.5% (SD 3.4%). Estimates for the log-normal distribution parameters are summarized in Table S1 (Supplementary Materials). The potential scale reduction factor \hat{R} —a diagnostic tool for the Bayesian sampling procedure—was always less than 1.01, indicating convergence of the MCMC routine. Estimated median RTs in the four different conditions are reported below in Table 1, along with expected mean, standard error and accuracy.

When the target appeared 50 ms after the face stimulus (early phase), an interaction effect between valence and target location was observed (e.s. = -0.048 , CI [-0.092 , -0.004]). This was due to slower RTs in OC trials compared to SC trials, for negative faces (e.s. of difference = 0.050 , CI [0.018 , 0.082]), while the evidence for an effect of target location on RTs for positive faces was not conclusive (e.s. of difference = 0.003 , CI [-0.030 , 0.036]).

On the other hand, when the target appeared 350 ms after the face stimulus (late phase), we did not observe a reliable interaction between valence and target location (e.s. = 0.046 , CI [-0.013 , 0.106]). However, as hypothesized, a simple effect of target location for negative faces was found (e.s. = -0.048 , CI [-0.090 , -0.005]) in the opposite direction compared to the 50 ms delay condition. These results are also shown graphically in Fig. 2.

Conclusions

With a short 50 ms lag between the presentation of emotional facial expressions and the response target, we observed faster response times to the target stimulus when this appeared perceptually closer to the face, rather than to the observer, for negative faces only. These findings support the hypothesis of an attentional priority of threatening expressions even when they are irrelevant for the task^{58–60} and, as a consequence, of a facilitation of target detection when the target is presented spatially close to face stimulus. Importantly, though, in the late phase (350 ms), the response to the target was facilitated when the latter was presented perceptually closer to the observer (rather than to the angry face)—providing evidence for a late attentional shift congruent with a potential retraction behavior from a threatening stimulus.

Experiment 2: fMRI study

To investigate the neural correlates of the late (350 ms) attentional withdrawal from threatening facial expressions observed in Experiment 1, we performed an fMRI study. The original stimulus presentation paradigm was adapted from the one in Experiment 1, using only the 350 ms lag between the presentation of the face stimulus and the response target.

Methods

Participants

Thirty-one right-handed young women (mean age: 26.1 ± 4.4 years; range 20–40 years) took part in the fMRI study. Since previous studies reported gender differences in empathy^{61,62}, only female volunteers were recruited. A history of psychiatric or neurological disease and current use of any psychoactive medications were considered as exclusion criteria, together with standard exclusion criteria from MR imaging. The sample size was chosen

	Condition	Median RT	Mean RT	SEM	Accuracy
Target 50 ms after	OC/pos	230.6	243.6	9.2	99.2%
	SC/pos	230.0	243.0	9.0	99.2%
	OC/neg	237.9	251.1	9.4	98.9%
	SC/neg	226.2	238.9	9.1	98.9%
Target 350 ms after	OC/pos	175.2	189.1	7.5	90.5%
	SC/pos	175.4	189.3	7.0	92.1%
	OC/neg	171.1	184.6	7.0	93.1%
	SC/neg	178.0	192.1	7.1	92.0%

Table 1. Expected values for median RT, mean RT and standard error (SEM) in the four conditions of interest in Experiment 1. Last column represents accuracy as the percentage of correct trials. Times are expressed in ms.

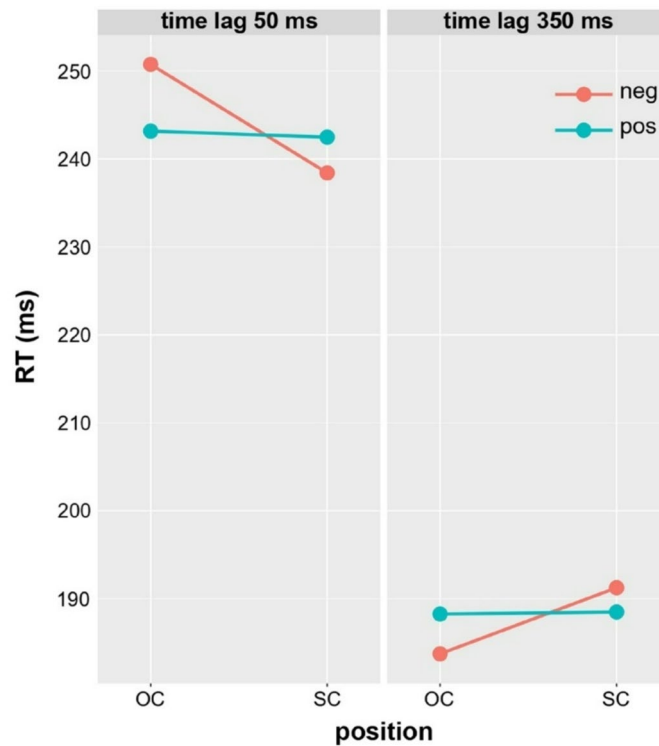


Fig. 2. Estimated average response times (RTs) for the four conditions of interest in Experiment 1. The left panel refers to the condition in which the target appears 50 ms after the face stimulus, the right panel to the condition in which the target appear 350 ms after the face stimulus. OC = observer-close, SC = stimulus-close, neg = negative faces, pos = positive faces.

based on common practice in task-based fMRI studies, where group sizes of approximately 20–30 are typical (median ≈ 28.5 in 2015⁶³). As reliable a priori power estimation in fMRI is rarely feasible due to unknown design-specific parameters^{64,65}, and given that significant results from studies of this scale remain valid under proper statistical control^{66,67}, we considered this number appropriate for our repeated-measures design. The study was conducted according to the 2013 version of the Declaration of Helsinki and had been approved by the local Ethics Committee (protocol number: CE 134/2014/SPER/AOUMO). All subjects gave their written informed consent to take part in the study.

Face stimuli

The same stimulus set used in Experiment 1 was employed, to which we added scrambled versions of the face images as control stimuli, with overall similar contrast and luminosity but devoid of recognizable facial features and thus of facial-expression related valence. This was done to allow evaluation of potential differences in spatial allocation of attention between meaningful and non-meaningful task-unrelated stimuli.

Procedure

A jittered event-related fMRI paradigm was employed. Participants received the task instructions before entering the scanner. Each volunteer performed 4 sessions of 84 trials each, for a total of 336 trials (112 with positive facial expressions, 112 with negative facial expressions, 112 with scrambled stimuli). Each trial lasted 2 s, with an intertrial interval ranging from 2 to 10 s. Emotional faces were visually presented via LCD goggles on the back wall of a corridor rendered in a 3D perspective. Participants were asked to respond as fast as possible to a target stimulus (a small red sphere, 100 ms), appearing 350 ms after the face in the SC (144 trials) or OC locations (144 trials), and gave their response by pressing a button on a response pad with the index finger of their right hand. A small number of foil trials was introduced: 32 faces (16 positive and 16 negative), plus 16 scrambled stimuli.

The `make_random_timing.py` routine of the AFNI software^{68,69} was used to create trial timing files for all stimulus categories in pseudo-random order. Briefly, 1000 sets of random timing files were simulated and the one with the highest statistical efficiency was selected. Two passive rest blocks were included at the beginning (duration 10–11.2 s) and at the end of each session (duration 22.3–27.2 s).

The E-Prime 3.0 software (Psychology Software Tools, Pittsburgh, PA) was used to deliver the stimuli via the NordicNeuroLab advanced VisualSystem fMRI goggles (<https://www.nordicneurolab.com>), and to collect behavioral responses. From the virtual point of view generated by the 3D-like presentation corresponding to the mid-point of the inferior border of the light blue transparent plane of Fig. 3, red-sphere targets closer to the face (SC condition, 61 pixels wide) subtended a visual angle of 10.5°, while targets closer to the observer (OC condition, 167 pixels wide) subtended a visual angle of 72°.

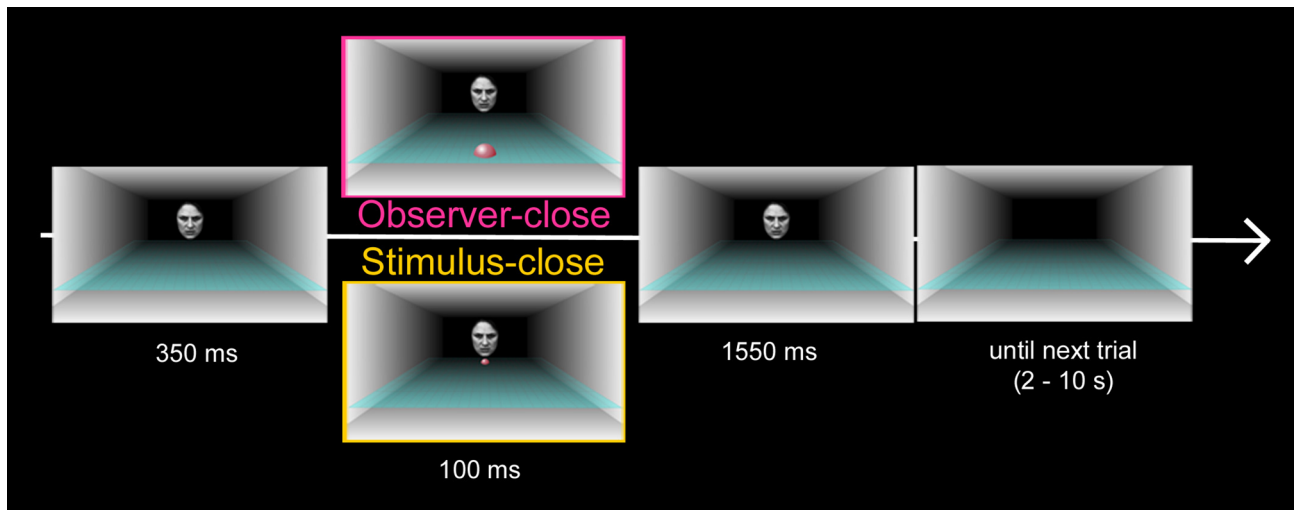


Fig. 3. Example screenshots of the distinct phases of a single trial and their corresponding durations. The face items in the figure have been adapted with permission from the Karolinska Directed Emotional Faces (KDEF; face ID: F01ANS⁷⁴).

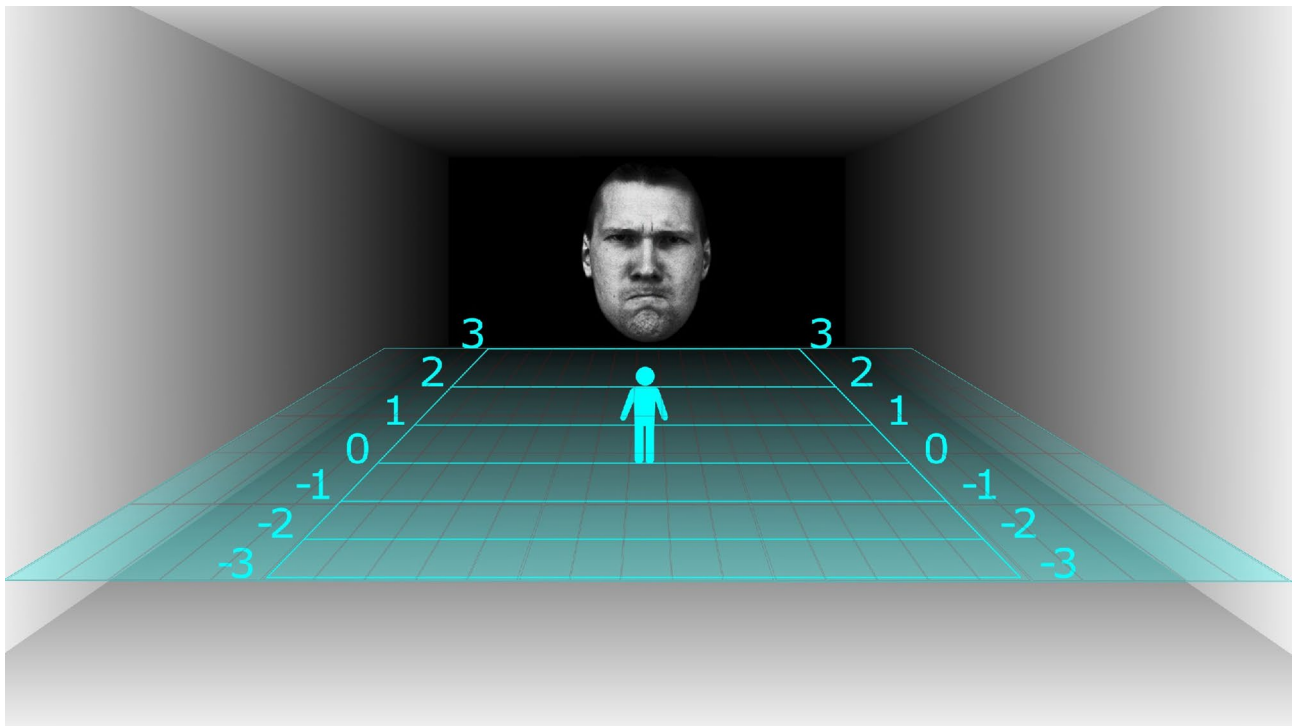


Fig. 4. 7-step staircase used for the post-scanning evaluation of subject-specific attraction and retraction feelings induced by face stimuli.

At the end of the scanning session, participants were asked to evaluate whether they felt—directionally, rather than emotionally—attracted to or repelled by the emotional facial expressions (positive and negative faces). Each stimulus was rated on a 7-point scale for each stimulus (-3 to $+3$), yielding a Subjective Attraction-Retraction (SAR) score for each face. SAR scores were collected only for positive and negative facial stimuli; no ratings were obtained for scrambled faces. SAR scores from $+1$ to $+3$ indicated feeling drawn to the face, scores from -1 to -3 feeling repelled by it, and a rating of zero indicated the absence of any such feeling. Stimuli in randomized order were shown on a laptop screen using the same 3D presentation settings used during scanning. A 7-step staircase was visualized on top of the floor of the 3D corridor (Fig. 4) with a stylized human figure at the middle point of the corridor (corresponding to a rating of zero). Then, all the experimental stimuli were also rated for their emotional valence using a 7-point scale, with positive values denoting positive valence. Finally,

participants completed the Interpersonal Reactivity Index (IRI)^{70,71} and the Behavioral Inhibition System/Behavioral Activation System scale (BIS/BAS)⁷². The IRI is a self-report rating index designed to measure personal empathy defined as the “reactions of one individual to the observed experiences of another”⁷³, and it consists of twenty-eight items and four sub-scales (Perspective Taking, Fantasy, Empathic Concern, Personal Distress). The BIS/BAS is a 24-items questionnaire comprising three BAS-related scales (Drive, Fun Seeking, Reward Responsiveness) and one BIS-related scale. We used the BIS/BAS scores as a validated assessment of subjective trait approach-avoidance tendencies, and the SAR scores as a phenomenological assessment of stimulus-specific approach-avoidance tendencies in the context of the present experiment.

Behavioral data analysis

Response times (RTs) for the six conditions of interest (positive, negative, or scrambled faces, with targets appearing in the OC or SC position) were assessed for each volunteer. Error and foil trials were excluded from the analysis. Responses faster than 150 ms were considered evidence that the volunteer responded to the onset of the face or scrambled stimulus, instead of responding to the target, and thus were also excluded from further analyses. We also set a maximum rate of false alarms (i.e., key presses occurring after face onset during foil trials) of 10% as an a priori exclusion criterion for participants. No participant exceeded this threshold (see Results).

The effects of valence and target location on response times were analyzed according to a shifted log-normal distribution, as in Experiment 1. One difference is that the factor coded as ‘valence’ in the previous model, was coded here as ‘face’ because of the additional presence of the scrambled stimuli. Thus, introducing the discrete variables v_1 , v_2 and t :

$$v_1 = \begin{cases} 0 & \text{if face is negative or scrambled} \\ 1 & \text{if face is positive} \end{cases} \quad v_2 = \begin{cases} 0 & \text{if face is negative or positive} \\ 1 & \text{if face is scrambled} \end{cases}$$

$$t = \begin{cases} 0 & \text{if target location is SC} \\ 1 & \text{if target location is OC} \end{cases}$$

the median RT in milliseconds can be expressed as $\mu = c + \exp(b_0 + v_1 b_1 + v_2 b_2 + t b_3 + t v_1 b_4 + t v_2 b_5)$ – see also Supplementary Materials. The hypothesis() function of the R package brms was used to provide effect size estimates (e.s.) and 95% credible intervals (CI) for the following contrasts: (1) scrambled versus facial stimuli (collapsed across valence), (2) the interaction between valence and target location, $(OC - SC)|_{\text{pos}} - (OC - SC)|_{\text{neg}}$, (3) the interaction between scrambled/negative stimuli and target location, $(OC - SC)|_{\text{scr}} - (OC - SC)|_{\text{neg}}$, and (iv) the simple effects of target location separately for positive, negative, and scrambled faces.

Analysis of subjective ratings and correlation with behavior, empathy and personality questionnaires

In order to verify the adequacy of the stimulus set for the purpose of the study, the across-subjects averages of reported valence and SAR scores for negative and positive faces were compared using two-sample t-tests.

Also, a subject-level correlation analysis between SAR and psychometric scores was performed using a Bayesian procedure, as follows. Differences in SAR scores between negative and positive face stimuli—an index of the individual sensitivity to the emotional valence of the faces for what concerns approach/avoidance feelings in the context of our experiment—were assessed against both the BIS/BAS and the IRI scores. The function correlationBF() of the R package BayesFactor was used⁷⁵, yielding an estimate of the posterior probability density over the correlation coefficient. This specific analysis was conducted on 26 out of 31 subjects, because IRI and BIS/BAS questionnaires were missing for five volunteers.

Finally, we tested the hypothesis that the individuals who showed the greater difference in SAR scores between negative and positive faces would be the ones exhibiting a larger behavioral effect for the interaction of valence and target location. Using the same Bayesian procedure, we performed an across-subjects correlation analysis between the magnitude of the above-mentioned interaction effect on RTs, on one side, and the average difference in SAR scores between negative and positive face stimuli, on the other.

fMRI data acquisition and preprocessing

fMRI data were acquired with a GE SIGNA Architect 3.0T MRI scanner and, for each volunteer, included 4 functional runs (gradient-echo echo-planar sequence, 46 2.7 mm-thick axial slices with a 0.3 mm gap, TR = 1500ms, in-plane resolution 3 × 3 mm), and a high-resolution T1-weighted anatomical image (344 sagittal slices, TR = 2184.9ms, TE = 3.09ms, voxel size = 1 × 1 × 1 mm). fMRI data were preprocessed and analyzed using the software package AFNI^{68,69}. Preprocessing included slice-timing and motion correction, warping of brain volumes to standard Montreal Neurological Institute (MNI) space, Gaussian spatial blurring (6 mm FWHM), and percent signal change scaling.

fMRI analysis

The effects of interest on BOLD activity were assessed by specifying and estimating a multiple-regression model for each subject. The event-related regressors representing the expected BOLD response included the following classes of stimuli: (1) positive SC trials, (2) positive OC trials, (3) negative SC trials, (4) negative OC trials, (5) scrambled SC trials, (6) scrambled OC trials, (7) positive foil trials, (8) negative foil trials, (9) scrambled foil trials. Each event was modeled starting at the time of appearance of the stimulus (positive, negative, or scrambled face) and lasting 2 s, followed by convolution with a gamma function accounting for the BOLD hemodynamic properties. Confounding covariates included the six motion parameters estimated during the motion correction phase, a 4th-order Legendre polynomial designed to model slow signal drifts unrelated to the experimental protocol, the first three principal components of the signal extracted from the lateral ventricle, and

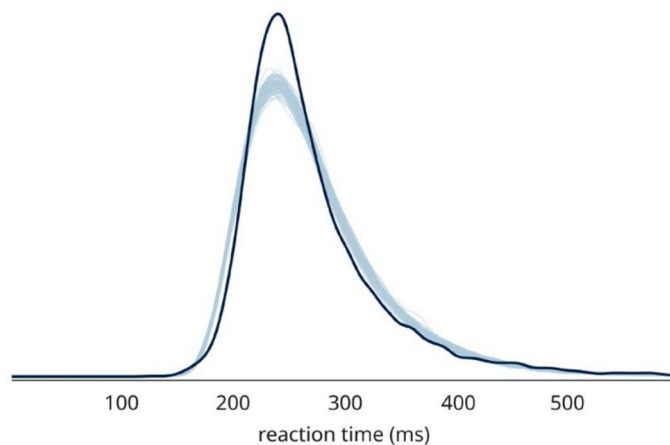


Fig. 5. Posterior predictive check for the analysis of RTs. The dark line indicates the original RT data distribution, while light blue lines represent simulated distributions based on 100 draws of parameter values from the posterior distribution. Good fitting performance is indicated by similar shape curves, peaking at similar RT values and with largely overlapping tails.

Condition	Median RT	Mean RT	SEM	Accuracy (%)
OC/pos	259.8	266.3	7.3	97.8
SC/pos	256.4	262.8	7.1	98.1
OC/neg	254.4	260.7	7.1	97.5
SC/neg	259.3	265.9	7.1	97.8
OC/scr	263.6	270.5	7.3	99.0
SC/scr	267.3	274.3	7.4	98.3

Table 2. Estimated values (ms) for median RT, mean RT and standard error of the mean (SEM) in the six conditions of interest. The last column represents response accuracy (percentage of correct responses).

a voxel-wise regressor computed by a locally-formed averaged signal from white-matter tissue^{76,77}. The model was fitted using a generalized least-square method allowing for the estimation of the temporal auto-correlation of the signal, using the AFNI routine 3dREMLfit.

In order to identify the neural correlates of spatial attention during positive and negative facial expression processing, we assessed the main effect of target location, (OC – SC), as well as the interaction of target location and valence, $(OC - SC)_{pos} - (OC - SC)_{neg}$. Group analyses were performed using a mixed-effect meta-analysis approach with the AFNI routine 3dMEMA⁷⁸, which models both within- and across-subject variability. The resulting group-level maps were thresholded using a combined single-voxel p value threshold and cluster extent criterion—as determined by the AFNI routine 3dClustSim via a Monte Carlo simulation of the cluster size distribution under the null hypothesis—to yield a family-wise error rate of $\alpha < 0.05$ ^{79,80}. Using the supra-threshold clusters of the obtained group statistical map for the *valence* \times *target location* effect as regions of interest (ROI), the average BOLD signal change was then extracted for each subject and for each experimental condition for further examination. With these data, an across-subjects Bayesian correlation analysis was also performed between the *valence* \times *target location* effect on BOLD signal change and the average difference in SAR scores between positive and negative face stimuli.

Results

Behavioral data

On average, participants reached an accuracy of 98.0% (SD = 2.3%). The mean false-alarm rate (i.e., response to foil trials) was 1.3% (SD = 1.9%), with a maximum of 5.6%, indicating that participants complied with instructions and did not respond in anticipation of the target. Estimates for the log-normal distribution parameters are summarized in Table S2 (Supplementary Materials). For each parameter, the potential scale reduction factor \hat{R} was less than 1.01, indicating convergence of the MCMC routine. The adequate fitting performance was further confirmed by the posterior predictive check displayed in Fig. 5, showing that the distributions generated by the sampling method, under the estimated values of the parameters, track well the actual data points.

Estimated median RTs, along with RT means, RT standard errors, and accuracy values, are reported in Table 2 and displayed graphically in Fig. 6. RTs were slower for scrambled compared to face stimuli, irrespective of valence (e.s. = 0.055, CI [0.032, 0.079]). We also observed an interaction between valence and target position, $(OC - SC)_{pos} - (OC - SC)_{neg}$ (e.s. = 0.058, CI [0.026, 0.091]). This resulted from faster response times for

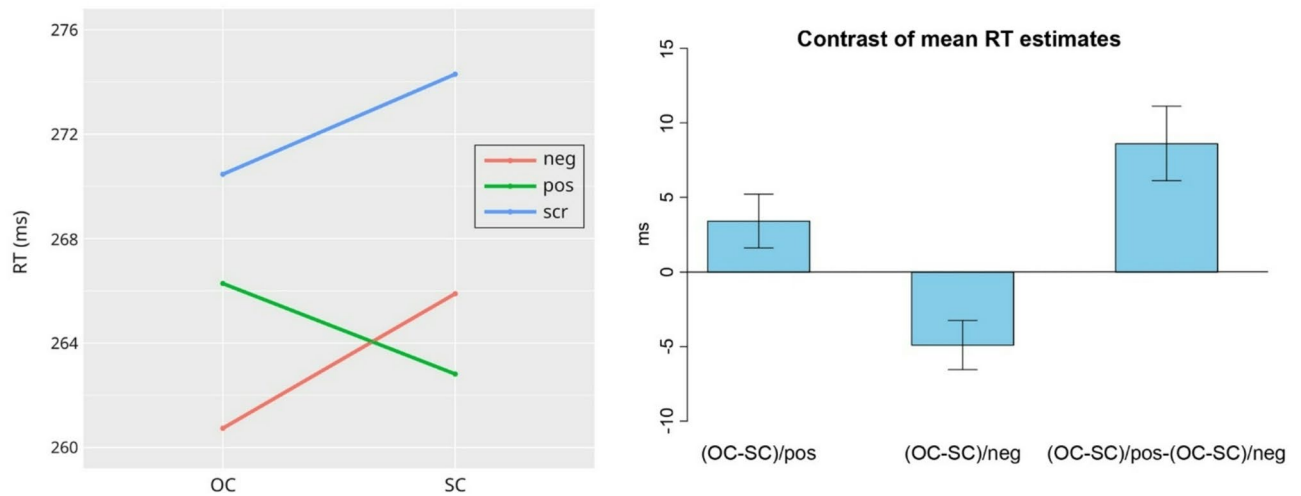


Fig. 6. Effects of valence and target location on response times. Left: interaction of face (pos, neg, scr) and target location (OC, SC). Right: selected contrasts of mean estimates, error bars indicate standard errors. OC = observer-close, SC = stimulus-close, *neg* = negative faces, *pos* = positive faces, *scr* = scrambled faces.

negative faces in the OC compared to the SC condition (e.s. = -0.035 , CI [-0.059 , -0.010]), while the evidence for a target location effect for positive faces (e.s. = 0.023 , CI [-0.002 , 0.049]) was not conclusive—although, numerically, response times for positive faces were *slower* in the OC compared to the SC condition. Finally, the evidence for the interaction between scrambled and negative stimuli with target location, $(OC-SC)|_{scr} - (OC-SC)|_{neg}$ (e.s. = 0.010 , CI [-0.023 , 0.044]) and for the effect of target location in scrambled faces (e.s. = -0.025 , CI [-0.053 , 0.003]) was not conclusive.

Subjective ratings and correlation with behavior, empathy and personality questionnaires

Reported valence and SAR scores were significantly different for negative and positive faces (average valence scores: positive faces = 2.05 , negative faces = -1.96 , $t(70) = -42.61$, $p < 0.001$; average SAR scores: positive faces = 1.20 , negative faces = -1.34 , $t(70) = -27.56$, $p < 0.001$). These results confirmed the validity of original classification of the stimulus set in terms of emotional valence for the purposes of our study.

Concerning the relationship between SAR scores and self-report questionnaires, only correlations with a coefficient whose 95% credible interval (CI) did not contain zero are reported here. The maximum a posteriori (MAP) estimate of the correlation between Empathic Concern (as measured by IRI) and the average difference in SAR scores between positive and negative stimuli was 0.50 (CI = [0.30 , 0.77]), with a Bayes Factor (BF) of 152.61 . The MAP estimate of the correlation between BIS score and the average difference in SAR scores between positive and negative stimuli was 0.40 (CI = [0.07 , 0.66]), with a BF of 5.85 .

The across-subjects correlation analysis between the magnitude of the *valence* \times *target location* effect on response times, on the one hand, and the average difference in SAR scores between positive and negative face stimuli, on the other hand, yielded a MAP estimate for the correlation coefficient of 0.40 (CI = [0.10 , 0.64]), with a BF of 7.19 . Figure 7 shows the posterior probability density for the correlation coefficient, along with the scatter plot of the two variables.

fMRI results

We first examined the main effect of target location on brain activation, and then we assessed the interaction of target location and valence to identify the regions showing an effect of target location with different magnitudes for positive and negative faces.

A significantly greater activation in the OC compared to the SC condition, independently of emotional valence, was exhibited by regions known to be involved in visual processing (primary and secondary visual areas), attentional deployment (intraparietal sulci and frontal eye fields)²⁶, arousal and salience detection (anterior insula and anterior cingulate)⁸¹, executive control (right dorso-lateral prefrontal cortex)⁸², and motor function (primary sensorimotor cortex, supplementary and pre-supplementary motor areas)—see Fig. 8 and Table S3 (Supplementary Materials).

The *valence* \times *target location* contrast yielded significant clusters (Fig. 9, top) in the bilateral occipital pole (mainly visual areas V1 and V2) and right middle occipital gyrus (V3 and V4), plus two small clusters in the left cerebellum (lobule VIIa) and thalamus/white matter, respectively (see also Table 3). The average BOLD signal percent changes in the occipital regions identified by the interaction contrast are shown in Fig. 9 (bottom), using the same graphical layout we used for RTs in Fig. 6 for comparison.

The across-subjects correlation analysis of the magnitude of the *valence* \times *target location* effect on BOLD signal change, on the one hand, and the average difference in SAR scores between positive and negative face stimuli, on the other, yielded a MAP estimate for the correlation coefficient of -0.32 (95% CI [-0.57 , -0.01]), with a BF of 2.75 . This amounted to a weak correlation between the two variables, where individuals with a

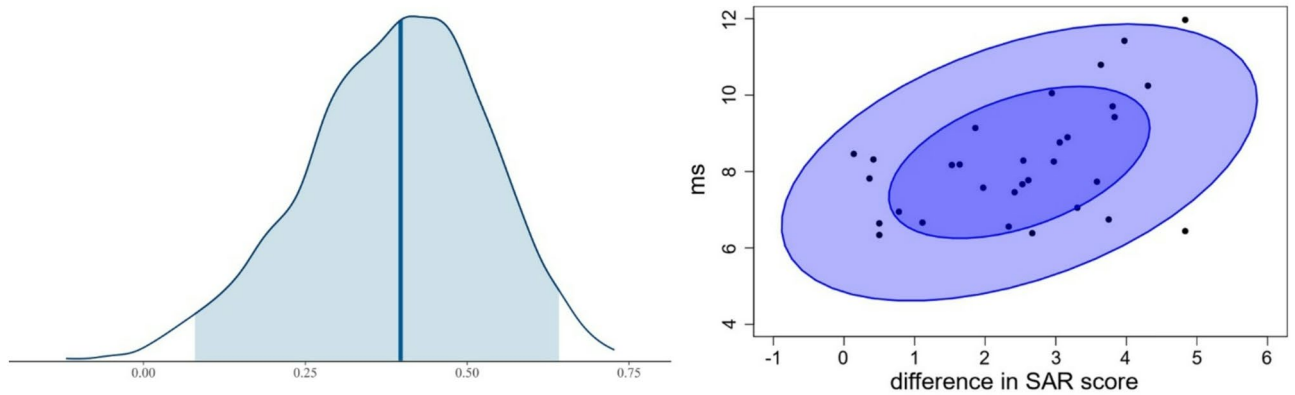


Fig. 7. Left panel: posterior distribution of the across-subjects correlation coefficient of the *valence* \times *target location* contrast on RTs and the difference in SAR scores between negative and positive stimuli. The shaded area indicates the 95% CI. Right panel: scatter plot of SAR score (*x*-axis) against the value of the *valence* \times *target location* effect on RTs (*y*-axis). The ellipses represent credible intervals: the dark-blue inner ellipse is the area containing 50% of the posterior distribution, and the pale-blue outer ellipse is the area containing 90% of the distribution.

Main effect of target location: OC > SC

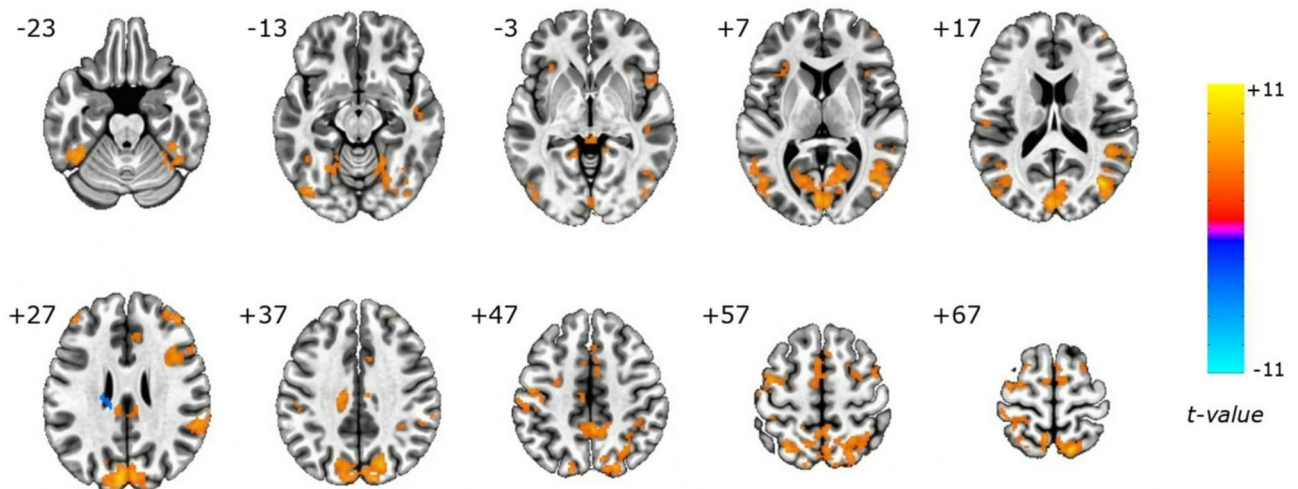


Fig. 8. Brain regions showing a significant main effect of target location, independently of stimulus valence. The *t*-map was obtained with a combined single-voxel threshold of $p < 0.001$ and a cluster extent threshold of 35 voxels, to give a whole-brain corrected $\alpha < 0.05$. The underlying anatomical image is the MNI template, and sections are displayed in neurological convention (left side of the brain is on the left), with numbers indicating the value of the MNI *z*-coordinate of the corresponding slice.

greater differential in SAR scores between positive and negative faces tended to exhibit a greater (negative) magnitude of the interaction effect on the BOLD signal change in the visual cortex.

Discussion

To the best of our knowledge, the present study is the first one to investigate a shift of spatial attention directionally congruent with a potential approach-avoidance behavior, conditional on the emotional valence of a stimulus. The behavioral data (both Experiment 1 and 2) support the hypothesis that the presentation of angry faces elicit an attentional movement with a valence-specific direction, characterized by an initial attentional capture, which is followed by an attentional retraction, away from the face stimulus and towards the observer. In Experiment 2, this behavioral effect correlated with how much each participant felt like ‘backing off’ from angry faces, a feeling that, in turn, was modulated by individual empathy and avoidance tendencies. The interaction between the emotional valence of the stimuli and the target location—the latter functioning as a probe of the current

Interaction of valence and target location

(OC-SC)/pos - (OC-SC)/neg

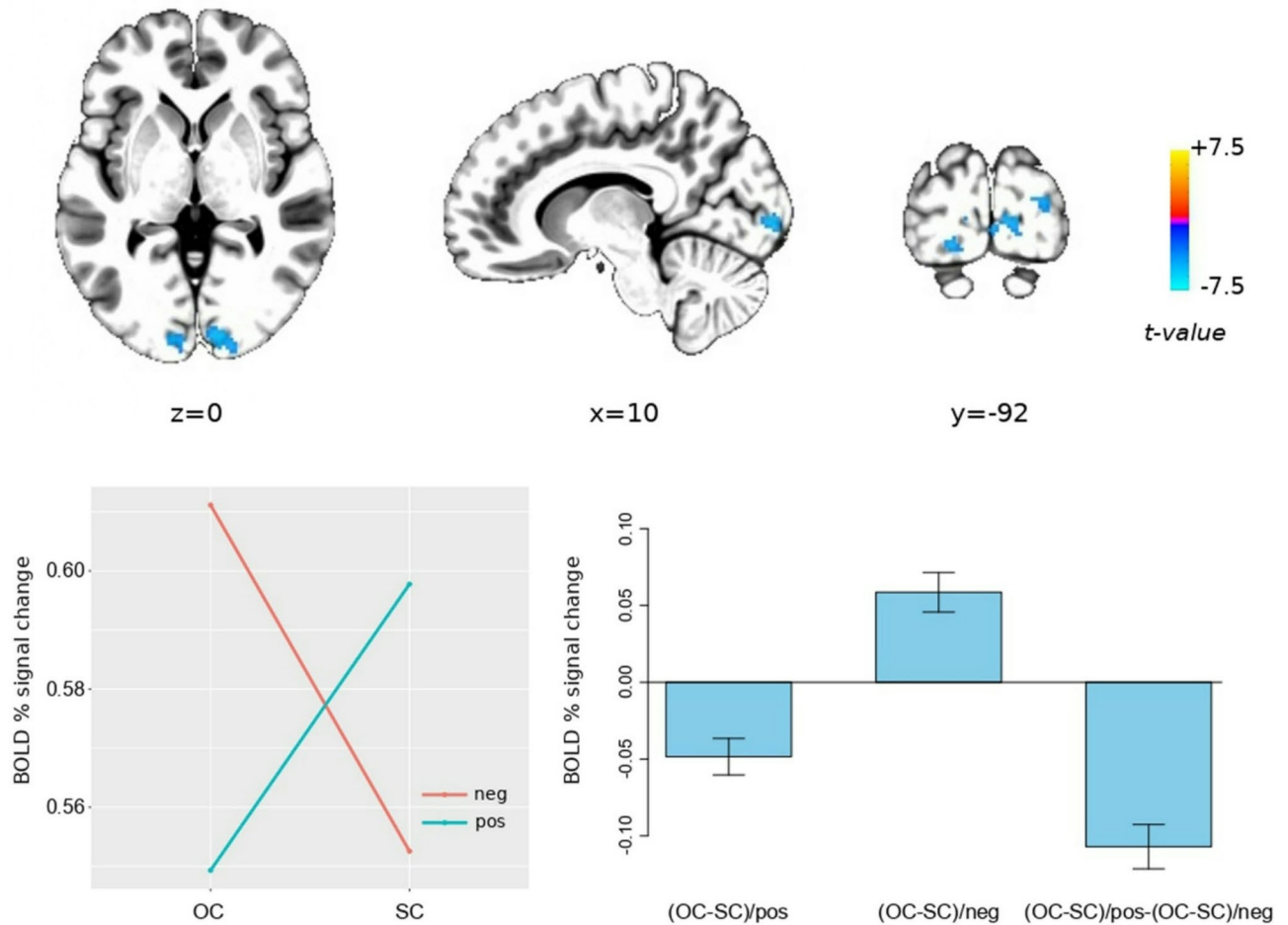


Fig. 9. Top: significant occipital brain activation clusters for the *valence × target location* effect. Bottom left: graphical display of the interaction. Bottom right: selected contrasts of parameter estimates for the average BOLD percent signal change in the occipital regions displayed on top, error bars indicate standard errors. OC = observer-close, SC = stimulus-close, *neg* = negative faces, *pos* = positive faces.

Brain region	k	X	y	z	t value
L V1/V2/V3	181	-13	-101	-9	-7.53
R V1	124	+3	-91	-3	-5.96
L Cerebellar lobule VIIa	101	-49	-67	-23	5.55
R V3/V4	53	+29	-89	+9	-5.81
LR Thalamus/white matter	35	+1	-5	+11	5.44

Table 3. Significant clusters for the *valence × target location* contrast. Cluster size (*k*) is in voxels, and the reported (MNI) coordinates (in mm) refer to each cluster’s statistical peak. L = left, R = right.

attentional locus—was associated with a significant change in the activation of early visual cortices, which correlated with the subjective feeling of attraction-repulsion. More specifically, activation in these regions was higher when the hypothesized attentional locus at the time of target presentation overlapped with the location of the latter.

The observed behavioral data provide new evidence for a differential time evolution of attention while processing happy versus angry facial expressions. Several ERPs studies have described the temporal dynamics of facial expression detection^{13,18}, showing attentional influences on early (P100, N170), mid-latency (P200, EPN), and late (P300, LPP) stages of processing (for a review, see⁶). Even though attention appears to be captured as

early as 100 ms from stimulus onset by a variety of threat stimuli, including angry faces^{83,84}, the duration of this ‘attentional grip’ by affective stimuli is still a matter of debate. In our first behavioral study, we investigated the effects on attention of an early (50ms) and a late (350ms) phase of emotional face processing. The results showed that response times to the target stimulus were faster when it appeared perceptually closer to the face, rather than to the observer, for negative faces only. This supported the hypothesis of an early automatic attentional capture by threatening faces^{58–60}, that would benefit target detection when the target appeared in a face’s nearby area of the visual field. Importantly, though, the response to the target was facilitated when the latter was presented perceptually closer to the observer (rather than to the negative face) *in the late phase only*—providing evidence for a late attentional shift elicited by angry faces, congruent with a potential retraction behavior. The fMRI study focused only on the late phase of face processing and confirmed the findings of the behavioral study, i.e., that for angry faces the response to the target was facilitated when the latter was presented perceptually closer to the observer. Crucially, an opposite effect was found for happy faces, showing a response facilitation for targets appearing perceptually closer to the face stimulus.

Previous studies have reported an early but fleeting capture of spatial attention by fearful faces^{20,22,85,86}; for instance, Torrence et al.⁸⁶ showed that fearful faces captured and held attention at times earlier than approximately 300 ms, while happy faces captured attention slightly later and held it longer. Other studies reported a more difficult or time-costly disengagement of attention from angry facial expressions⁸⁷, especially for high trait anxious individuals^{88,89}, even though, in social anxiety, attentional disengagement from threatening stimuli has been reported from 300ms post stimulus onset⁹⁰. More recently, shorter latencies to disengage attention from angry faces have been reported, but using only 200 and 500 ms latencies as experimental values⁸⁹. The more pronounced EEG contingent negative variation (CNV) recorded by Schutter et al.⁸⁷ over parietal sites for angry facial expressions (within a 425–525 ms time window) was considered by the authors as a possible neural correlate of sustained attention, but also of action preparation. Thus, the attentional disengagement from angry faces appears to be around 300–500 ms post-stimulus onset—a range consistent with the value chosen in our study and the present findings—although the precise timing of the process is likely to depend also on the specific characteristics of the emotional stimulus and experimental protocol (e.g., passive viewing or response task). Interestingly, scrambled faces elicited slower overall response times than either positive or negative faces, and the SC–OC effect was in the same direction as that observed for negative faces, although it did not reach statistical significance. We interpret this pattern as potentially reflecting the perceptual ambiguity or incongruity of scrambled stimuli. The difficulty in recognizing such stimuli may increase cognitive interference, leading to slower responses, while at the same time endowing them with a mildly aversive quality. In general, ambiguous or poorly identifiable stimuli tend to be treated with caution, which could explain their functional similarity to negative faces. As we had not anticipated this effect, we did not collect SAR measures for this class of stimuli, and our interpretation must therefore remain tentative.

An effect of compatibility between affect and action tendencies³⁵ has been shown using facial expressions as stimuli in behavioral tasks, where pushing away aversive faces and pulling appetitive stimuli towards oneself were considered automatic responses associated with shorter reaction times^{41,42}. However, there is also evidence in the literature for the elicitation of appetitive and/or aversive behaviors by both positive and negative stimuli, in contrast with the hypothesis of an automatic and direct compatibility effect between affect and action tendency. For instance, facial expression of anger has been linked to avoidance, but also to approach tendencies^{43–45}, whereas happiness may elicit avoidance reaction when expressed by out-group members⁴⁶. The automaticity of the association between emotion processing and approach-avoidance tendencies was questioned also by two recent meta-analyses^{47,48}, concluding that the link between affective stimuli and muscle movements is indirect, malleable and can be contextually modulated. The first one⁴⁷ found a small-to-medium facilitation effect by positive/negative stimuli on approach/avoidance tendencies, an absence of the effect for implicit evaluation tasks, and a modulation of the results by the reference frame (self vs. object). The second meta-analysis⁴⁸ investigated the facilitation of arm flexion/extension movements as a function of stimulus valence, showing that the context (affective and goal-related cues) can modulate the kinematics of arm movements, with a significantly larger compatibility effect for emotional faces (compared to words or scenes) in general, and for angry expressions in particular.

In our study, we aimed to investigate the shift of attention during emotional face processing in the direction linking the observer to the face stimulus, predicting an effect of compatibility between attentional shift and approach-avoidance tendencies elicited by emotional stimuli. The findings proved consistent with the hypothesis of a late attentional retraction from angry facial expressions and towards the observer. A specific influence of negative emotional content on the spatial allocation of selective attention has been previously hypothesized. For instance, a recent study²⁰ investigated the distribution of spatial attention triggered by looming fearful faces, as opposed to neutral and happy faces, appearing in the near or far peripersonal space⁹¹, demonstrating a shift of attention to the periphery when a fearful face was presented closer to the body. The authors interpreted this effect as related to a peripheral search for potential sources of threat, whose presence was hinted to by the fearful facial expression. From the perspective of peripersonal space, the stimulus-close condition in our experimental design is akin to a far-space condition, whereas the observer-close condition can be linked to the near space of the participant. Unlike facial expressions of fear, however, an angry face such as the ones employed in our study coincides with the potential source of threat, so that when it appears in the far space, a late attentional retraction to the near space may indicate the individual’s preparation to execute a distancing movement from the threat. This interpretation is in agreement with our hypothesis of a congruence between attentional shift and approach-avoidance tendencies, but also with the premotor theory of attention, which posits that spatial attention is underwritten by neural circuits that mediate the conversion of spatial information into action⁹². When motor programs for acting in space are prepared, but not immediately executed, this translates into an

increased motor readiness and a facilitation of stimulus processing for the spatial location aimed to by the motor plan, effectively bringing about a close association of attentional and premotor processes.

Interestingly, in our study, the spatial shift of attention suggested by reaction times was correlated with the participants' feeling of attraction or aversion elicited by positive and negative faces; specifically, the facilitation of behavioral responses when the target was presented perceptually closer to the observer for the angry faces, or closer to the face stimulus for the happy ones, was greater for those individuals who showed a stronger subjective feeling of aversion from the negative faces or attraction to the positive ones, as measured by the SAR scores. In turn, this subjective feeling of attraction and repulsion was correlated to the individual score for empathic concern measured by IRI⁷⁰—as well as, more cogently, to the individual behavioral avoidance tendency as measured by the BIS scale. Thus, the present findings provide further evidence for a mechanism of emotion-related attentional avoidance as proposed by Becker & Detweiler-Bedell⁹³, who acquired eye-tracking data during passive viewing of a series of emotional faces, demonstrating that the rapid processing of the potential threat of fearful and angry faces induces an active avoidance of negative faces (within 300ms of display onset). This mechanism may play a role in maintaining maladaptive response patterns in anxious individuals, as proposed by the so-called 'vigilance-avoidance' hypothesis^{49,50}.

Our fMRI data showed a significant interaction between emotional valence and target location in the activity of early visual cortices. This finding is in line with previous studies showing that early visual areas can display both an emotional attention and a spatial attention effect^{22–24}. Several brain regions, including category-specific areas (i.e., the fusiform face area for facial expressions), but also non-specific ones, such as the early sensory cortices or fronto-parietal attentional circuits, have been shown to exhibit an increased response to emotional versus neutral stimuli⁹⁴. This modulation is considered the hallmark of the so-called emotional attention effect⁹⁵, which implicates a lower threshold for the detection of emotional stimuli and a preferential processing of affective content⁹⁶. The early sensory cortices play a critical role for the modality-specific representations of valence⁹⁷—as their activity can be used to decode the valence of visual stimuli⁹⁸—and they are involved in constructing modality-specific affective experience⁹⁹. Previous studies have also shown that the activity of the early visual areas, such as V1, V2 and V3, can be modulated by intense emotional stimuli^{21,100}, emotional attention^{23,101,102}, and spatial attention^{103–105}. Indeed, emotionally salient content can influence the visual cortex via automatic bottom-up processes^{23,26}, but also through a top-down attentional modulation by the fronto-parietal network implicated in selecting visual locations^{27–29}. A recent fMRI study examined the effects of emotional pictures and a spatial cue, hinting either to the left or to the right visual field, on the activity of visual areas, showing an enhanced response in V4, intraparietal sulcus and lateral occipital cortex for the spatial cue, and in V3, V4 and lateral occipital cortex for the emotional pictures²⁴. This is in line with the data in the present study, where the activation of early visual areas and V4 was significantly modulated by both emotional and spatial attention. Furthermore, the greater response of these areas when the target appeared closer to the face stimulus for the happy faces, and closer to the observer for the angry faces (Fig. 9, bottom left), together with the faster reaction times in these conditions (Fig. 6, left), supports the hypothesized spatial congruence at the time of behavioral response between the locus of attention and the target location, compatible with a late attentional retraction triggered by angry faces. Finally, the finding that participants who felt more repelled by the angry faces or attracted to the happy ones showed a larger interaction effect of valence and target location on the activation of visual cortices, provides further evidence for the hypothesized congruence between spatial attention dynamics and individual approach-avoidance tendencies.

Limitations and future directions

While the results of this study support the hypothesized influence of emotional facial expressions on the dynamics of spatial attention—and a role of early visual cortices in mediating this effect—we acknowledge a number of limitations that should be addressed in further experimental work. First, we used facial expressions as emotional visual stimuli, with only two types of emotions (i.e., happy and angry); a more general applicability of the findings could be investigated by using different facial expressions or generic emotional pictures of objects and scenes, such as those in the IAPS dataset¹⁰⁶. Second, because SAR scores were not collected for scrambled stimuli, correlation analyses between response times and SAR scores were limited to positive and negative faces. This omission constitutes a limitation of the study, as it precludes a more precise interpretation of the response time pattern observed for scrambled faces. Third, the use of electrophysiological techniques with high temporal resolution (ERPs), recorded simultaneously with fMRI, could clarify with greater precision the temporal dynamics of the attentional retraction elicited by angry faces and its association with the activity of early visual cortices. Moreover, the use of a single stimulus-onset asynchrony (350 ms) as a fixed delay makes target onsets predictable and thus potentially reduce across-trial variability in response times, independent of experimental conditions. To discourage expectancy-driven responses we inserted 48 foil trials in which no target appeared and, indeed, the very low false-alarm rate in these foil trials argues against the presence of a relevant anticipatory component. However, in order to investigate the evolution of the spatial allocation of attention in the context of emotional stimuli at a finer time scale, a targeted study using a larger number of temporal lags should be performed. Finally, to investigate covert attention shifts, participants were instructed to maintain their gaze on the face stimulus while responding to the target. Future studies could gain important insights by relaxing this constraint and incorporating eye-tracking, thereby enhancing the ecological validity of the findings.

Data availability

The dataset used and analysed for the current study is available from the corresponding author on reasonable request.

Received: 28 November 2023; Accepted: 22 October 2025

Published online: 24 November 2025

References

- Compton, R. J. The interface between emotion and attention: a review of evidence from psychology and neuroscience. *Behav. Cogn. Neurosci. Rev.* **2**, 115–129 (2003).
- Taylor, J. G. & Fragopanagos, N. F. The interaction of attention and emotion. *Neural Netw.* **18**, 353–369 (2005).
- Palermo, R. & Rhodes, G. Are you always on my mind? A review of how face perception and attention interact. *Neuropsychologia* **45**, 75–92 (2007).
- Brosch, T., Scherer, K., Grandjean, D. & Sander, D. The impact of emotion on perception, attention, memory, and decision-making. *Swiss Med. Wkly.* **143**, w13786. (2013).
- Dolcos, F. et al. Neural correlates of emotion-attention interactions: from perception, learning, and memory to social cognition, individual differences, and training interventions. *Neurosci. Biobehav. Rev.* **108**, 559–601 (2020).
- Schindler, S. & Bublatzky, F. Attention and emotion: an integrative review of emotional face processing as a function of attention. *Cortex* **130**, 362–386 (2020).
- Goodhew, S. C. & Edwards, M. Don't look now! Emotion-induced blindness: the interplay between emotion and attention. *Atten. Percept. Psychophys.* **84**, 2741–2761 (2022).
- Broadbent, D. E. *Perception and Communication* (Pergamon, 1958). <https://doi.org/10.1037/10037-000>
- Treisman, A. M. Selective attention in man. *Br. Med. Bull.* **20**, 12–16 (1964).
- Chelazzi, L., Della Libera, C., Sani, I. & Santandrea, E. Neural basis of visual selective attention. *Wiley Interdiscip. Rev. Cogn. Sci.* **2**, 392–407 (2011).
- Fiebelkorn, I. C. & Kastner, S. Functional specialization in the attention network. *Annu. Rev. Psychol.* **71**, 221–249 (2020).
- Eimer, M., Holmes, A. & McGlone, F. P. The role of Spatial attention in the processing of facial expression: an ERP study of rapid brain responses to six basic emotions. *Cogn. Affect. Behav. Neurosci.* **3**, 97–110 (2003).
- Luo, W., Feng, W., He, W., Wang, N. Y. & Luo, Y. J. Three stages of facial expression processing: ERP study with rapid serial visual presentation. *Neuroimage* **49**, 1857–1867 (2010).
- Batty, M. & Taylor, M. J. Early processing of the six basic facial emotional expressions. *Brain Res. Cogn. Brain Res.* **17**, 613–620 (2003).
- Eger, E., Jedynak, A., Iwaki, T. & Skrandies, W. Rapid extraction of emotional expression: evidence from evoked potential fields during brief presentation of face stimuli. *Neuropsychologia* **41**, 808–817 (2003).
- Miyoshi, M., Katayama, J. & Morotomi, T. Face-specific N170 component is modulated by facial expressional change. *Neuroreport* **15**, 911–914 (2004).
- Caharel, S., Courtney, N., Bernard, C., Lalonde, R. & Rebai, M. Familiarity and emotional expression influence an early stage of face processing: an electrophysiological study. *Brain Cogn.* **59**, 96–100 (2005).
- Sun, L., Ren, J. & He, W. Neural correlates of facial expression processing during a detection task: an ERP study. *PLoS One* **12**, e0174016 (2017).
- Holmes, A., Vuilleumier, P. & Eimer, M. The processing of emotional facial expression is gated by spatial attention: evidence from event-related brain potentials. *Brain Res. Cogn. Brain Res.* **16**, 174–184 (2003).
- Ellena, G., Starita, F., Haggard, P. & Ládavas, E. The Spatial logic of fear. *Cognition* **203**, 104336 (2020).
- Ballotta, D., Maramotti, R., Borelli, E., Lui, F. & Pagnoni, G. Neural correlates of emotional Valence for faces and words. *Front. Psychol.* **14**, 1055054 (2023).
- Pourtois, G. & Vuilleumier, P. Dynamics of emotional effects on Spatial attention in the human visual cortex. In *Understanding Emotions* 67–91 (Elsevier, (2006).
- Vuilleumier, P. & Driver, J. Modulation of visual processing by attention and emotion: windows on causal interactions between human brain regions. *Philos. Trans. R Soc. Lond. B Biol. Sci.* **362**, 837–855 (2007).
- Bi, T. et al. Modulations of emotional attention and spatial attention on human visual cortical activities. *Psychol. Res. Behav. Manag.* **12**, 375–384 (2019).
- Corbetta, M. & Shulman, G. L. Human cortical mechanisms of visual attention during orienting and search. *Philos. Trans. R Soc. Lond. B Biol. Sci.* **353**, 1353–1362 (1998).
- Corbetta, M. & Shulman, G. L. Control of goal-directed and stimulus-driven attention in the brain. *Nat. Rev. Neurosci.* **3**, 201–215 (2002).
- Desimone, R. & Duncan, J. Neural mechanisms of selective visual attention. *Annu. Rev. Neurosci.* **18**, 193–222 (1995).
- Martínez, A. et al. Involvement of striate and extrastriate visual cortical areas in Spatial attention. *Nat. Neurosci.* **2**, 364–369 (1999).
- Squire, R. F., Noudoost, B., Schafer, R. J. & Moore, T. Prefrontal contributions to visual selective attention. *Annu. Rev. Neurosci.* **36**, 451–466 (2013).
- Blair, R. J. R. Facial expressions, their communicatory functions and neuro-cognitive substrates. *Philos. Trans. R Soc. Lond. B Biol. Sci.* **358**, 561–572 (2003).
- Horstmann, G. What do facial expressions convey: feeling states, behavioral intentions, or actions requests? *Emotion* **3**, 150–166 (2003).
- Parkinson, B. Do facial movements express emotions or communicate motives? *Pers. Soc. Psychol. Rev.* **9**, 278–311 (2005).
- Frijda, N. H. *The Emotions* (Cambridge University Press, 1986).
- Lang, P. J., Bradley, M. M. & Cuthbert, B. N. Emotion, attention, and the startle reflex. *Psychol. Rev.* **97**, 377–395 (1990).
- Solarz, A. K. Latency of instrumental responses as a function of compatibility with the meaning of eliciting verbal signs. *J. Exp. Psychol.* **59**, 239–245 (1960).
- Chen, M. & Bargh, J. A. Consequences of automatic evaluation: immediate behavioral predispositions to approach or avoid the stimulus. *Pers. Soc. Psychol. Bull.* **25**, 215–224 (1999).
- De Houwer, J., Crombez, G., Baeyens, F. & Hermans, D. On the generality of the affective Simon effect. *Cogn. Emot.* **15**, 189–206 (2001).
- Markman, A. B. & Brendl, C. M. Constraining theories of embodied cognition. *Psychol. Sci.* **16**, 6–10 (2005).
- Rinck, M. & Becker, E. S. Approach and avoidance in fear of spiders. *J. Behav. Ther. Exp. Psychiatry* **38**, 105–120 (2007).
- Saraiva, A. C., Schüür, F. & Bestmann, S. Emotional valence and contextual affordances flexibly shape approach-avoidance movements. *Front. Psychol.* **4**, 933 (2013).
- Cacioppo, J. T., Priester, J. R. & Berntson, G. G. Rudimentary determinants of attitudes: II. Arm flexion and extension have differential effects on attitudes. *J. Pers. Soc. Psychol.* **65**, 5–17 (1993).
- Marsh, A. A., Ambady, N. & Kleck, R. E. The effects of fear and anger facial expressions on approach- and avoidance-related behaviors. *Emotion* **5**, 119–124 (2005).
- Rotteveel, M. & Phaf, R. H. Automatic affective evaluation does not automatically predispose for arm flexion and extension. *Emotion* **4**, 156–172 (2004).
- Carver, C. S. & Harmon-Jones, E. Anger is an approach-related affect: evidence and implications. *Psychol. Bull.* **135**, 183–204 (2009).

45. Krieglmeyer, R. & Deutsch, R. Approach does not equal approach. *Soc. Psychol. Pers. Sci.* **4**, 607–614 (2013).
46. Paulus, A. & Wentura, D. Threatening joy: approach and avoidance reactions to emotions are influenced by the group membership of the expresser. *Cogn. Emot.* **28**, 656–677 (2014).
47. Phaf, R. H., Mohr, S. E., Rotteveel, M. & Wicherts, J. M. Approach, avoidance, and affect: a meta-analysis of approach-avoidance tendencies in manual reaction time tasks. *Front. Psychol.* **5**, 378 (2014).
48. Laham, S. M., Kashima, Y., Dix, J. & Wheeler, M. A meta-analysis of the facilitation of arm flexion and extension movements as a function of stimulus valence. *Cogn. Emot.* **29**, 1069–1090 (2015).
49. Mogg, K., Mathews, A. & Weinman, J. Memory bias in clinical anxiety. *J. Abnorm. Psychol.* **96**, 94–98 (1987).
50. Mathews, A. Why worry? The cognitive function of anxiety. *Behav. Res. Ther.* **28**, 455–468 (1990).
51. Radke, S. et al. Oxytocin reduces amygdala responses during threat approach. *Psychoneuroendocrinology* **79**, 160–166 (2017).
52. Olivier, J. & Norberg, M. Positively skewed data: revisiting the Box-Cox power transformation. *Int. J. Psychol. Res.* **3**, 69–78 (2010).
53. van der Linden, W. J. A lognormal model for response times on test items. *J. Educ. Behav. Stat.* **31**, 181–204 (2006).
54. Rouder, J. N. Are unshifted distributional models appropriate for response time? *Psychometrika* **70**, 377–381 (2005).
55. Cohen, A. C. Three-parameter estimation. In *Lognormal Distributions: Theory and Applications* (eds Crow, E. & Shimizu, K.) 113–137 (Routledge, (1987).
56. Ranger, J., Kuhn, J. T. & Ortner, T. M. Modeling responses and response times in tests with the hierarchical model and the three-parameter lognormal distribution. *Educ. Psychol. Meas.* **80**, 1059–1089 (2020).
57. Bürkner, P. C. Brms: an R package for bayesian multilevel models using Stan. *J. Stat. Soft* **80**, 1–28 (2017).
58. Hodsoll, S., Viding, E. & Lavie, N. Attentional capture by irrelevant emotional distractor faces. *Emotion* **11**, 346–353 (2011).
59. Burra, N., Barras, C., Coll, S. Y. & Kerzel, D. Electrophysiological evidence for attentional capture by irrelevant angry facial expressions. *Biol. Psychol.* **120**, 69–80 (2016).
60. Huang, S. L., Chang, Y. C. & Chen, Y. J. Task-irrelevant angry faces capture attention in visual search while modulated by resources. *Emotion* **11**, 544–552 (2011).
61. Decety, J. & Jackson, P. L. The functional architecture of human empathy. *Behav. Cogn. Neurosci. Rev.* **3**, 71–100 (2004).
62. Singer, T. The neuronal basis and ontogeny of empathy and mind reading: review of literature and implications for future research. *Neurosci. Biobehav. Rev.* **30**, 855–863 (2006).
63. Poldrack, R. A. et al. Scanning the horizon: towards transparent and reproducible neuroimaging research. *Nat. Rev. Neurosci.* **18**, 115–126 (2017).
64. Desmond, J. E. & Glover, G. H. Estimating sample size in functional MRI (fMRI) neuroimaging studies: statistical power analyses. *J. Neurosci. Methods* **118**, 115–128 (2002).
65. Mumford, J. A. & Nichols, T. E. Power calculation for group fMRI studies accounting for arbitrary design and Temporal autocorrelation. *NeuroImage* **39**, 261–268 (2008).
66. Friston, K. Ten ironic rules for non-statistical reviewers. *NeuroImage* **61**, 1300–1310 (2012).
67. Friston, K. Sample size and the fallacies of classical inference. *NeuroImage* **81**, 503–504 (2013).
68. Cox, R. W. AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Comput. Biomed. Res.* **29**, 162–173 (1996).
69. Cox, R. W. & Hyde, J. S. Software tools for analysis and visualization of fMRI data. *NMR Biomed.* **10**, 171–178 (1997).
70. Davis, M. A. Multidimensional approach to individual differences in empathy. *JSAS Cat. Sel. Doc. Psychol.* **10**, 85 (1980).
71. Albiero, P., Ingoglia, S. & Coco, A. L. Interpersonal reactivity index—Italian version. *PsyCTESTS Dataset* <https://doi.org/10.1037/t81465-000> (2006).
72. Carver, C. S. & White, T. L. Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: the BIS/BAS scales. *J. Pers. Soc. Psychol.* **67**, 319–333 (1994).
73. Davis, M. H. Measuring individual differences in empathy: evidence for a multidimensional approach. *J. Personal. Soc. Psychol.* **44**, 113–126 (1983).
74. Lundqvist, D., Flykt, A., Öhman, A. & Karolinska, D. Emotional faces. <https://doi.org/10.1037/t27732-000> (1998).
75. Morey, R. D. & Rouder, J. N. *BayesFactor: Computation of Bayes Factors for Common Designs* (2023).
76. Jo, H. J., Saad, Z. S., Simmons, W. K., Milbury, L. A. & Cox, R. W. Mapping sources of correlation in resting state FMRI, with artifact detection and removal. *Neuroimage* **52**, 571–582 (2010).
77. Jo, H. J. et al. Effective preprocessing procedures virtually eliminate distance-dependent motion artifacts in resting state FMRI. *J. Appl. Math.* (2013).
78. Chen, G., Saad, Z. S., Nath, A. R., Beauchamp, M. S. & Cox, R. W. FMRI group analysis combining effect estimates and their variances. *NeuroImage* **60**, 747–765 (2012).
79. Cox, R. W., Chen, G., Glen, D. R., Reynolds, R. C. & Taylor, P. A. fMRI clustering and false-positive rates. *Proc. Natl. Acad. Sci. USA* **114**, E3370–E3371 (2017).
80. Cox, R. W., Chen, G., Glen, D. R., Reynolds, R. C. & Taylor, P. A. FMRI clustering in AFNI: false-positive rates redux. *Brain Connect.* **7**, 152–171 (2017).
81. Menon, V. & Uddin, L. Q. Saliency, switching, attention and control: a network model of Insula function. *Brain Struct. Funct.* **214**, 655–667 (2010).
82. Panikratova, Y. R. et al. Functional connectivity of the dorsolateral prefrontal cortex contributes to different components of executive functions. *Int. J. Psychophysiol.* **151**, 70–79 (2020).
83. Cooper, R. M. & Langton, S. R. H. Attentional bias to angry faces using the dot-probe task? It depends when you look for it. *Behav. Res. Ther.* **44**, 1321–1329 (2006).
84. Koster, E. H. W., Verschuere, B., Crombez, G. & Van Damme, S. Time-course of attention for threatening pictures in high and low trait anxiety. *Behav. Res. Ther.* **43**, 1087–1098 (2005).
85. Carlson, J. M. & Reinke, K. S. Attending to the fear in your eyes: facilitated orienting and delayed disengagement. *Cogn. Emot.* **28**, 1398–1406 (2014).
86. Torrence, R. D., Wylie, E. & Carlson, J. M. The time-course for the capture and hold of visuospatial attention by fearful and happy faces. *J. Nonverbal Behav.* **41**, 139–153 (2017).
87. Schutter, D. J. L. G., de Haan, E. H. F. & van Honk, J. Functionally dissociated aspects in anterior and posterior electrocortical processing of facial threat. *Int. J. Psychophysiol.* **53**, 29–36 (2004).
88. Fox, E., Russo, R. & Dutton, K. Attentional bias for threat: evidence for delayed disengagement from emotional faces. *Cogn. Emot.* **16**, 355–379 (2002).
89. Leleu, V., Douilliez, C. & Rusinek, S. Difficulty in disengaging attention from threatening facial expressions in anxiety: a new approach in terms of benefits. *J. Behav. Ther. Exp. Psychiatry* **45**, 203–207 (2014).
90. Moriya, J. & Tanno, Y. The time course of attentional disengagement from angry faces in social anxiety. *J. Behav. Ther. Exp. Psychiatry* **42**, 122–128 (2011).
91. Rizzolatti, G., Fadiga, L., Fogassi, L. & Gallese, V. The space around us. *Science* **277**, 190–191 (1997).
92. Craighero, L., Rizzolatti, G., Itti, L., Rees, G. & Tsotsos, J. K. CHAPTER 31—The Premotor Theory of Attention. In *Neurobiology of Attention* 181–186 (Academic Press, Burlington, 2005). <https://doi.org/10.1016/B978-012375731-9/50035-5>
93. Becker, M. W. & Detweiler-Bedell, B. Early detection and avoidance of threatening faces during passive viewing. *Q. J. Exp. Psychol. (Hove)* **62**, 1257–1264 (2009).

94. Lindquist, K. A., Wager, T. D., Kober, H., Bliss-Moreau, E. & Barrett, L. F. The brain basis of emotion: a meta-analytic review. *Behav. Brain Sci.* **35**, 121–143 (2012).
95. Vuilleumier, P. How brains beware: neural mechanisms of emotional attention. *Trends Cogn. Sci.* **9**, 585–594 (2005).
96. Pourtois, G., Schettino, A. & Vuilleumier, P. Brain mechanisms for emotional influences on perception and attention: what is magic and what is not. *Biol. Psychol.* **92**, 492–512 (2013).
97. Miskovic, V. & Anderson, A. K. Modality general and modality specific coding of hedonic valence. *Curr. Opin. Behav. Sci.* **19**, 91–97 (2018).
98. Shinkareva, S. V. et al. Representations of modality-specific affective processing for visual and auditory stimuli derived from functional magnetic resonance imaging data. *Hum. Brain Mapp.* **35**, 3558–3568 (2014).
99. Satpute, A. B. et al. Involvement of sensory regions in affective experience: a meta-analysis. *Front. Psychol.* **6**, 1860 (2015).
100. Bradley, M. M. et al. Activation of the visual cortex in motivated attention. *Behav. Neurosci.* **117**, 369–380 (2003).
101. Keil, A., Moratti, S., Sabatinelli, D., Bradley, M. M. & Lang, P. J. Additive effects of emotional content and spatial selective attention on electrocortical facilitation. *Cereb. Cortex* **15**, 1187–1197 (2005).
102. Gomez, A. et al. Emotion modulates the effects of endogenous attention on retinotopic visual processing. *Neuroimage* **57**, 1542–1551 (2011).
103. Gandhi, S. P., Heeger, D. J. & Boynton, G. M. Spatial attention affects brain activity in human primary visual cortex. *Proc. Natl. Acad. Sci. USA* **96**, 3314–3319 (1999).
104. Somers, D. C., Dale, A. M., Seiffert, A. E. & Tootell, R. B. Functional MRI reveals spatially specific attentional modulation in human primary visual cortex. *Proc. Natl. Acad. Sci. USA* **96**, 1663–1668 (1999).
105. Watanabe, T. et al. Attention-regulated activity in human primary visual cortex. *J. Neurophysiol.* **79**, 2218–2221 (1998).
106. Lang, P. J., Bradley, M. M. & Cuthbert, B. N. *International Affective Picture System (IAPS): Affective Ratings of Pictures and Instruction Manual. Technical Report A-8* (2008).

Author contributions

D.B., E.B., F.L., and G.P. contributed to the design of the experiment. D.B., R.M., E.B., and G.P. contributed to data collection. D.B., R.M., E.B., and G.P. contributed to data analysis. All authors contributed to the interpretation of the results and manuscript writing.

Funding

This study was awarded by the Ministero dell’Istruzione, dell’Università e della Ricerca (Project title: “The Good and the Bad of Sensory Experience: Understanding the Impact of Emotionally Charged Stimuli on Cognition and Behavior, and the Brain’s Mechanisms to Cope with Them” Grant number: CUP E94I19000630005).

Declarations

Competing interests

The authors declare that they have no competing interest.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-025-25609-w>.

Correspondence and requests for materials should be addressed to D.B.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025